

RESEARCH DEPARTMENT

SUBJECTIVE JUDGMENTS ON BRIGHTNESS

Report No. T.043
Serial No. 1954/3

Report written by:

M. Gilbert, Ph.D., B.Sc., D.I.C.

W. Proctor Wilson

(W. Proctor Wilson)

REPORT NO. T.043

SUBJECTIVE JUDGMENTS ON BRIGHTNESS

<u>Section</u>	<u>Title</u>	<u>Page</u>
1	INTRODUCTION - - - - -	1
2	EXPERIMENTAL - - - - -	2
3	NOMENCLATURE - - - - -	4
4	RESULTS - - - - -	4
	TABLE 1 - - - - -	4
	TABLE 2 - - - - -	5
	TABLE 3 - - - - -	6
	CONCLUSIONS - - - - -	7
	REFERENCES - - - - -	8
	FIG. 1 - - - - -	Facing page 2
	FIG. 2 - - - - -	" " 6
	FIG. 3 - - - - -	" " 6

CONFIDENTIAL

Research Department

REPORT NO. T.043

January, 1954.

Serial No. 1954/3

Figs. 1 to 3.

SUBJECTIVE JUDGMENTS ON BRIGHTNESS

1. INTRODUCTION

(1)(2)
Experiments have been conducted in the field of acoustics in order to determine whether it is possible for people to judge with any accuracy that one sound is twice as loud as another standard sound, or half as loud as the standard sound. It was found that people can be persuaded to make judgments of "twice as loud" and their results appear to be consistent within reasonable limits. It was concluded that one sound needs to be approximately ten times the intensity of another sound (or have ten times the energy) for it to seem twice as loud, to a statistically average listener.

Symposia on subjective judgments, such as that organised by the Building Research Station in October 1950, have given rise to considerable argument on the possibility of making a judgment other than in respect of the equality of two sensations. The judgment may involve any of the five sensations of touch, smell, hearing, vision or taste. The fundamental point in question is whether a person can say only that two sensations are equal or that one is greater than the other, or whether he can say that one sensation is (say) twice as great as another or is (say) two sensation units more than the other.

In the field of vision, if people were found to agree with reasonable consistency on the amount of light that looked twice as bright as another light, the criterion could be of practical value. For example, a knowledge of how much more energy is needed to make a picture appear twice as bright would indicate the extent to which it might be worth while increasing the light energy of the picture appearing on the face of the television tube.

The multi-criterion method of obtaining subjective judgments would be of very great value in this particular context. Hopkinson, found that by asking subjects to make a set of inter-related judgments, such as:-

The light is (a) just perceptibly brighter

(b) definitely brighter

(c) very much brighter

(d) intensely more bright.

than the control source, the complete set of judgments was far more self-consistent than if a single criterion were asked for. Such a set of judgments would provide an arbitrary scale of magnitude of sensation against light energy values as measured by a meter. The scale would be of value in the example of television tube brightness mentioned above, but it could not be compared with any other arbitrary scale of sensation values. Since the sensation "twice as much" has already been studied in the field of acoustics, it was, therefore, decided to assess this sensation-ratio in the field of vision in order to be able to interpret quantitatively a direct comparison between two different types of sensation.

2. EXPERIMENTAL

A diagram of the apparatus is shown in Fig. 1.

In order to obtain two independent patches of light for comparison and to provide adjustment of brightness, two separate but similar light paths were used, each including a car headlamp bulb as a point source of light. Light from the filament was diffused into a uniformly bright patch using a piece of ground glass. This patch was used to flash a condenser lens so that it appeared as a uniformly bright disk when viewed on axis. A fixed and a rotating polaroid were placed in front of the lens, and an arbitrary scale engraved around the first polaroid and a pointer on the second polaroid enabled the relative positions of the pair to be assessed whilst observations were being made. The scale was calibrated in terms of field brightness, measured on axis through the polaroids. Each of the two fields was given a nominal calibration with the aid of an S.E.I. photometer. These nominal calibrations were used to set the two fields at theoretically equal brightness levels and the equality was checked visually. Minor adjustments were made when necessary, and the check thus enabled a higher degree of accuracy to be achieved in calibrating the apparatus.

A right-angle prism placed in front of each field reflected half the circular patch of light, so that two semi-circular patches could be viewed simultaneously side-by-side along a path at right angles to the axis of the original light paths. Monocular vision was necessary to enable the lens images to be viewed centrally, and thus appear as

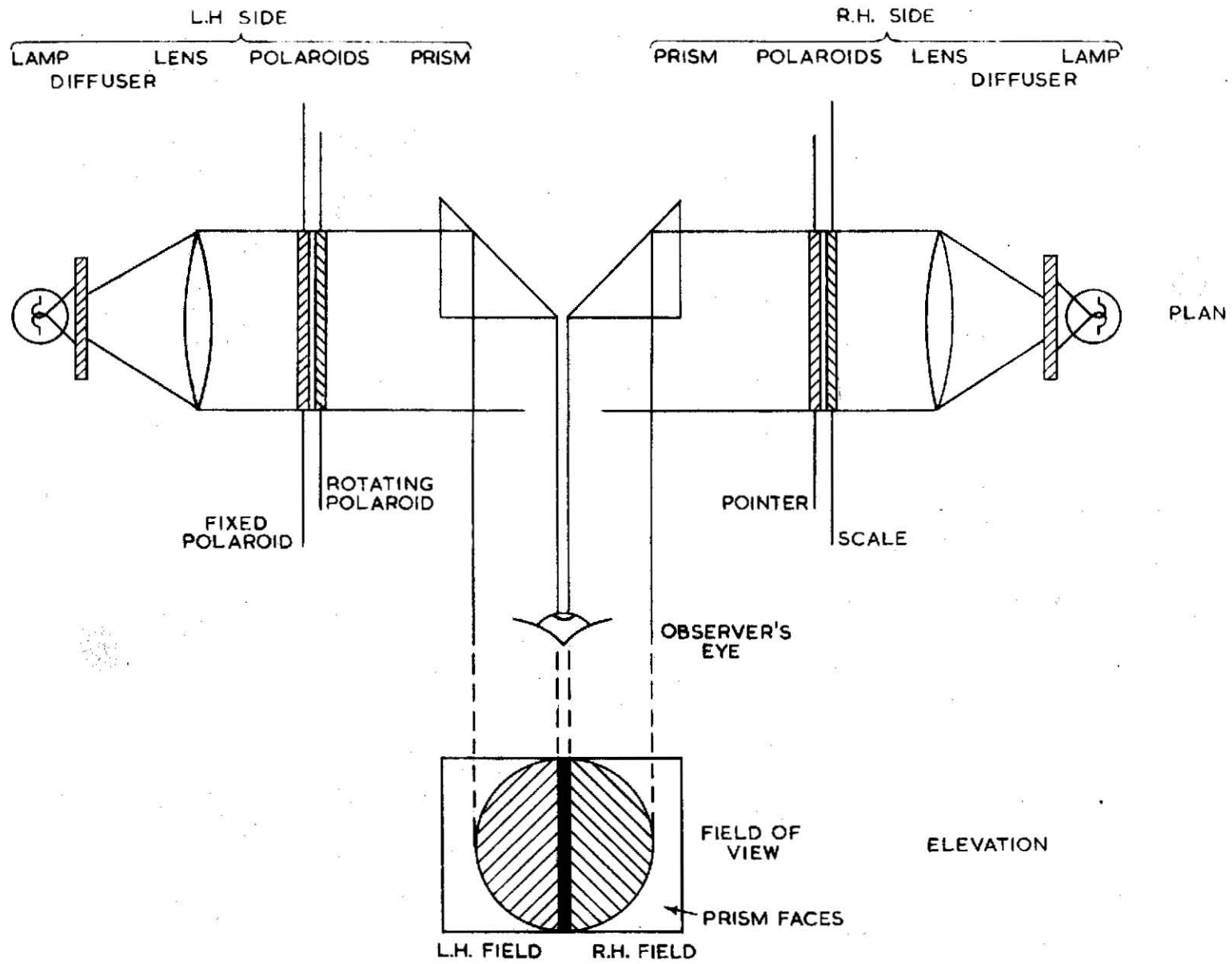


FIG.1
DIAGRAM OF APPARATUS

uniformly bright, and a chin rest was provided in order to eliminate undue strain in maintaining the eye in a fairly steady position about 13 inches away from the centre of the bipartite field. The angle subtended at the eye by the field was $5\frac{1}{2}^{\circ}$.

The brightness of the surroundings to this central circular field was maintained at a constant level of about 0.3 to 0.4 ft. lamberts. The brightness values corresponding with "twice-as-bright" and "half-as-bright" were investigated at five levels of brightness of the central fields:- 1.5, 0.0, 1.0 and 1.5 log (ft. lamberts).

The subjective tests were conducted in the following manner:

One of these levels was selected at random and for a first group of tests it was set on the right half of the field, and a test subject was requested to rotate the left-hand moving polaroid until the two half-fields appeared equally bright. This was repeated six times in order to obtain an average reading. The observer then went on to make six judgments of "twice-as-bright" and then six of "half-as-bright". A second group of tests was then undertaken with the pre-determined brightness levels set on the left half of the field and adjustments were made by the observer on the right half of the field.

Brightness matching was included as well as the "twice-" and "half-as-bright" criteria in order to ascertain the variations or statistical "spread" of the results obtained with the present apparatus when a familiar type of measurement was made. This enables the performance of observers on this type of apparatus to be compared with the performance achieved on the apparatus of other workers, and from this comparison the relationships of "twice-" and "half-as-bright" can be placed alongside known standards.

The left half and the right half of the field were used in turn as the pre-determined brightness levels, because it might be possible for an observer to have a bias in one direction. For example, an observer might always make the adjustable field slightly brighter than the standard field when endeavouring to match the two, but as each side in turn is made the adjustable field, this error is eliminated. It was also hoped that this procedure would eliminate the "coefficient of involuntary mendacity" that may always be present in experiments of this nature.

The eleven subjects who made the measurements received the idea of "twice-as-bright" with mixed emotions. No one really felt it was a reasonable request, and many thought it to be impossible or completely meaningless. All were persuaded to try it, however, without being given any indication of what might be meant or required by the phrase, and

after one or two readings had been taken they seemed to gain confidence that such measurements were possible.

3. NOMENCLATURE

In the results to be discussed, the individual observations, which were repeated six times, have been termed "attempts" and only the mean values of any set of six attempts are quoted. The mean values are termed "readings". The mean values of any group or groups of readings have been called "averages". Thus, an "average" of eleven readings is calculated from data arising from sixty-six attempts.

4. RESULTS

At first the brightness of the right-hand field was adjusted to a nominal value and that of the left-hand field visually matched to it. The scale readings obtained in this way from the left-hand field at each of the five nominal brightness levels were assumed to correspond to the exact values of the nominal brightness given on the right-hand field. When, on the other hand, the left-hand field was made the standard and the brightness of the right-hand matched to it, the right-hand brightness readings obtained were noted. Table 1 shows these readings and also the standard deviations of each group of twenty-two readings.

TABLE 1

BRIGHTNESS MATCHING

(Averages for 11 subjects, 2 reading each, 6 attempts per reading)

L.H. "Readings"	$\bar{1}.50$	0.00	0.50	1.00	1.50	log ft.-lamberts
R.H. "	$\bar{1}.52$	$\bar{1}.96$	0.47	0.99	1.55	" " "
L.H.-R.H. "	-0.02	+0.04	+0.03	+0.01	-0.05	" units
L.H. Standard Deviation:	0.04	0.04	0.10	0.03	0.04	" "
R.H. " "	0.05	0.06	0.07	0.05	0.06	" "

* Calibrated values.

These figures prove that both sets of values are statistically of the same distribution, because the differences between the means of the "readings" are less than their standard deviations. The right- and left-side

variables have substantially the same average value at each of the five brightness levels. This would be expected, and it indicates the degree of reliability of the calibration and of the subjects making the observations. It was decided, therefore, that it would be justifiable to average the "readings" whether they were obtained by varying the left-hand field or the right-hand field. The first three lines of Table 2 appertain to the "averages" obtained in this way, at each of the nominal brightness levels investigated. The standard deviations of each group of "twice-as-bright" and "half-as-bright" estimations are given, and they may be compared with the standard deviations of each group of brightness matches.

The objective brightness ratios resulting from the subjective criteria of "two times" and "half times" are shown in this table. They are given in logarithmic units, the actual ratios being the antilogarithms of the numbers given in lines (4), (5), (8) and (9) of the table.

TABLE 2

(Averages for 11 subjects, 2 readings each, 6 attempts per reading)

(1) Brightness Level:	$\bar{1}.50$	0.00	0.50	1.00	1.50	log ft. lamberts		
(2) "Average" of L.H. and R.H. "Readings":	$\bar{1}.51$	$\bar{1}.98$	0.49	1.00	1.53	"	"	"
(3) Standard Deviation:	0.05	0.07	0.07	0.05	0.09	"	units	
(4) Two-Times Ratio:	0.64	0.46	0.40	0.67	0.57	"	"	
(5) Half-Times Ratio:	-0.40	-0.58	-0.52	-0.54	-0.67	"	"	
(6) Standard Deviation (Two-Times Ratio):	0.10	0.23	0.29	0.25	0.26	"	"	
(7) Standard Deviation (Half-Times Ratio):	0.25	0.23	0.22	0.39	0.25	"	"	
(8) Geometric Mean of All Two-Times Ratios:						0.55 log units		
(9) Geometric Mean of All Half-Times Ratios:						-0.54	"	"

There is a significant difference between the highest and the lowest ratio for either criterion, but on the other hand there does not seem to be any definite trend or obvious relationship between the ratios and the

brightness levels, as may be seen in Fig. 2. There is no steady increase or decrease of the ratio with brightness.

Fig. 3 shows the distributions of choices made by the subjects of values for the ratios "twice" and "half", and it can be seen that there is no significant difference between them. It is, in fact, highly probable that they are the same. Their standard deviations at each brightness level are also similar, so that it appears to be equally difficult to judge "half-as-bright" as "twice-as-bright", and brightness is additive or subtractive by similar psychological mechanisms within the limits of this experiment. The standard deviations, in log units, are about five times those of the matching readings.

Although the readings of all the subjects have been averaged, it does not necessarily follow that one person's idea of "twice" is the same as another person's, and Table 3 has been constructed to show the averages for each subject, assuming that all the brightness levels may be taken together for this purpose.

TABLE 3

(Averages for 5 brightness levels, 2 readings at each level,
6 attempts per reading)

Subject:	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>	<u>I</u>	<u>J</u>	<u>K</u>		
Twice Ratio:	0.66	0.35	0.65	0.20	0.37	0.50	1.03	0.61	0.40	0.53	0.74	log units	
Half "	<u>0.72</u>	<u>0.27</u>	<u>0.54</u>	<u>0.34</u>	<u>0.37</u>	<u>0.59</u>	<u>0.83</u>	<u>0.57</u>	<u>0.46</u>	<u>0.36</u>	<u>0.92</u>	"	"
Mean:	0.69	0.31	0.60	0.27	0.37	0.55	0.93	0.59	0.43	0.45	0.83	"	"
<hr/>													
Standard Deviations:													
Twice Ratio:	0.27	0.11	0.38	0.11	0.11	0.22	0.22	0.11	0.11	0.11	0.16	"	"
Half "	0.18	0.09	0.16	0.17	0.17	0.18	0.21	0.15	0.15	0.15	0.18	"	"

The standard deviations of the readings of most of the subjects are lower than the standard deviations shown in Table 2, which seems to indicate that it is more justifiable to average the different brightness-level values than to average all the subjects' values. The difference between the means of the subjects with the highest and lowest readings (subject D and subject G) is highly significant. They very definitely have different criteria for the idea of "twice" and "half", but it is noted that each

100-17
4.6.53
APP'D m.
REPORT TO 43
This drawing specification is the property of the British Broadcasting Corporation and may not be reproduced or disclosed to a third party in any form without the written permission of the Corporation.
BBC
DS/44/P

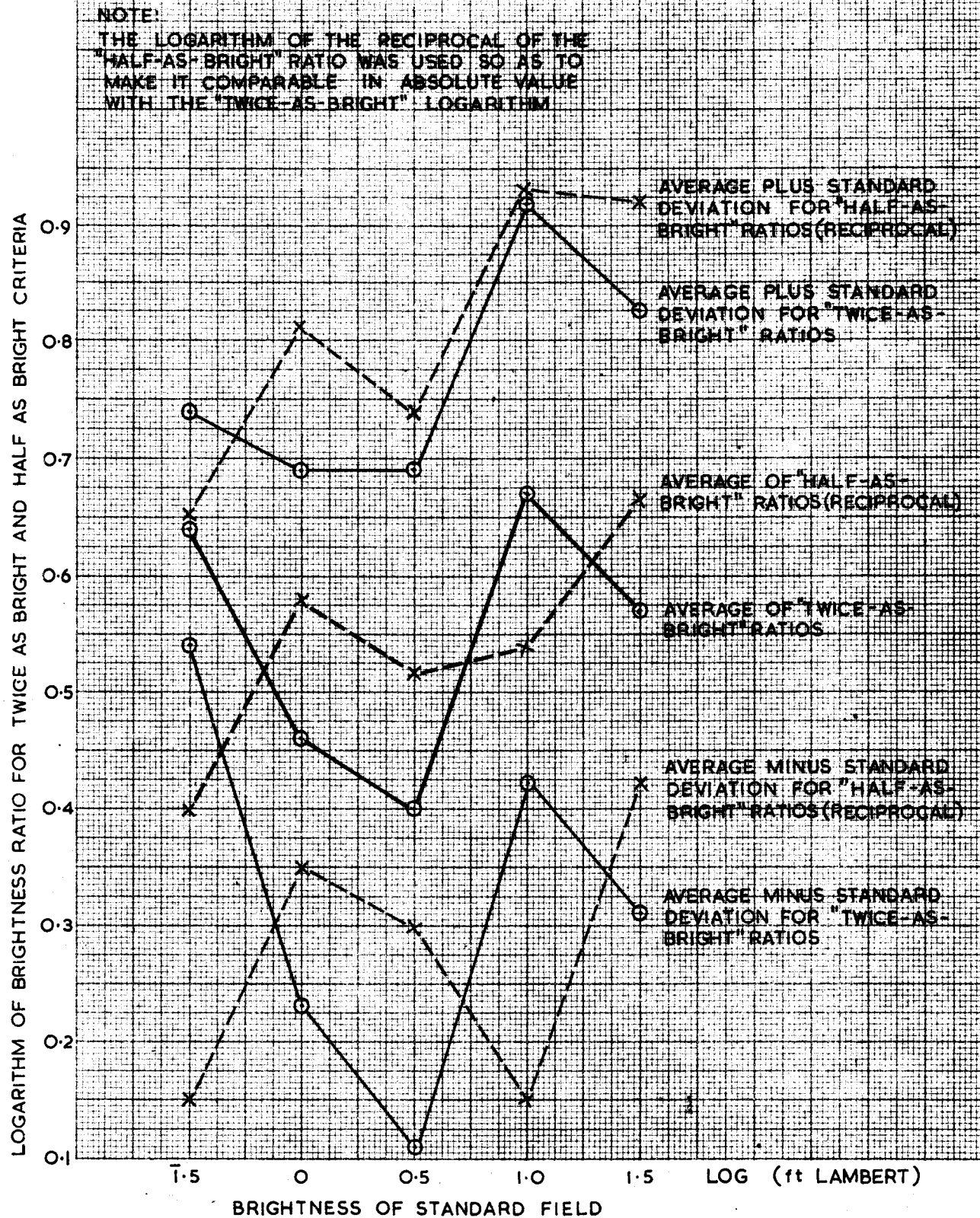
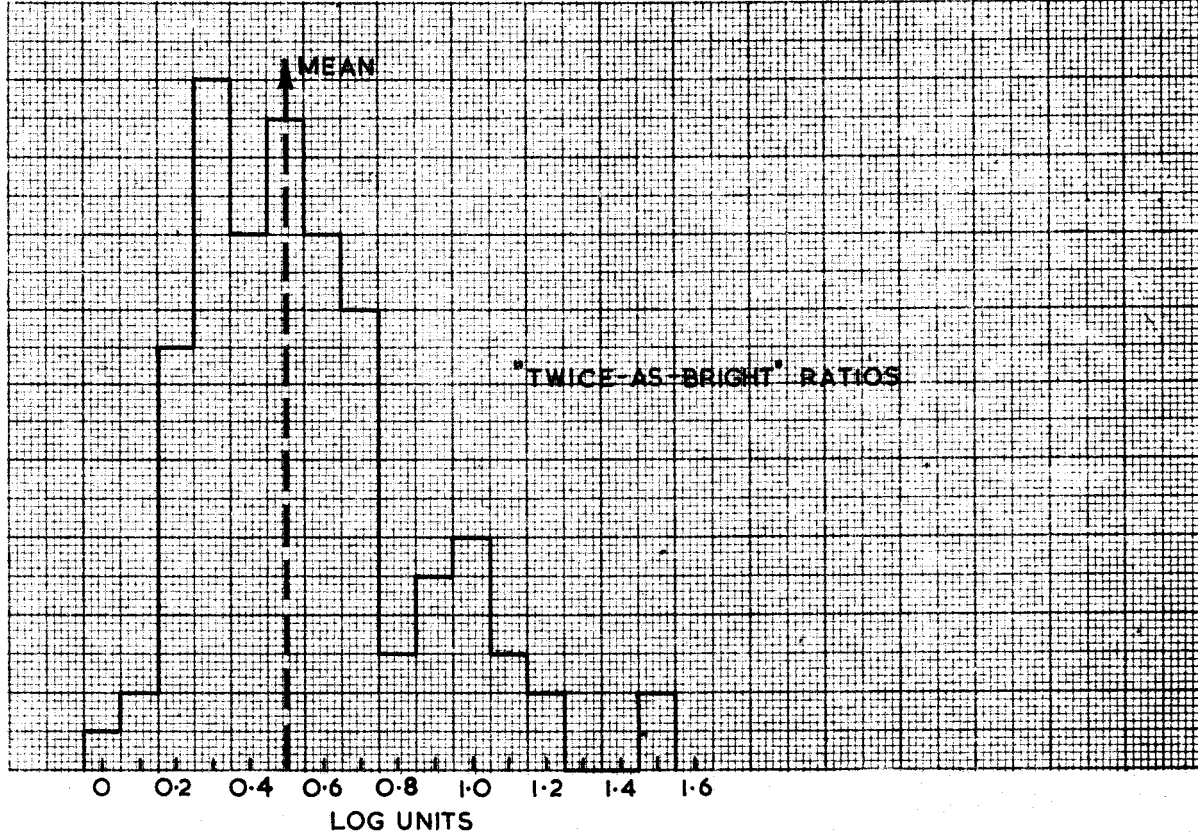


FIG. 2

VARIATION OF AVERAGE LOGARITHM OF "TWICE-AS-BRIGHT" AND "HALF-AS-BRIGHT" RATIO WITH BRIGHTNESS LEVEL.

NUMBER CHOOSING ABSCISSA VALUE



NUMBER CHOOSING ABSCISSA VALUE

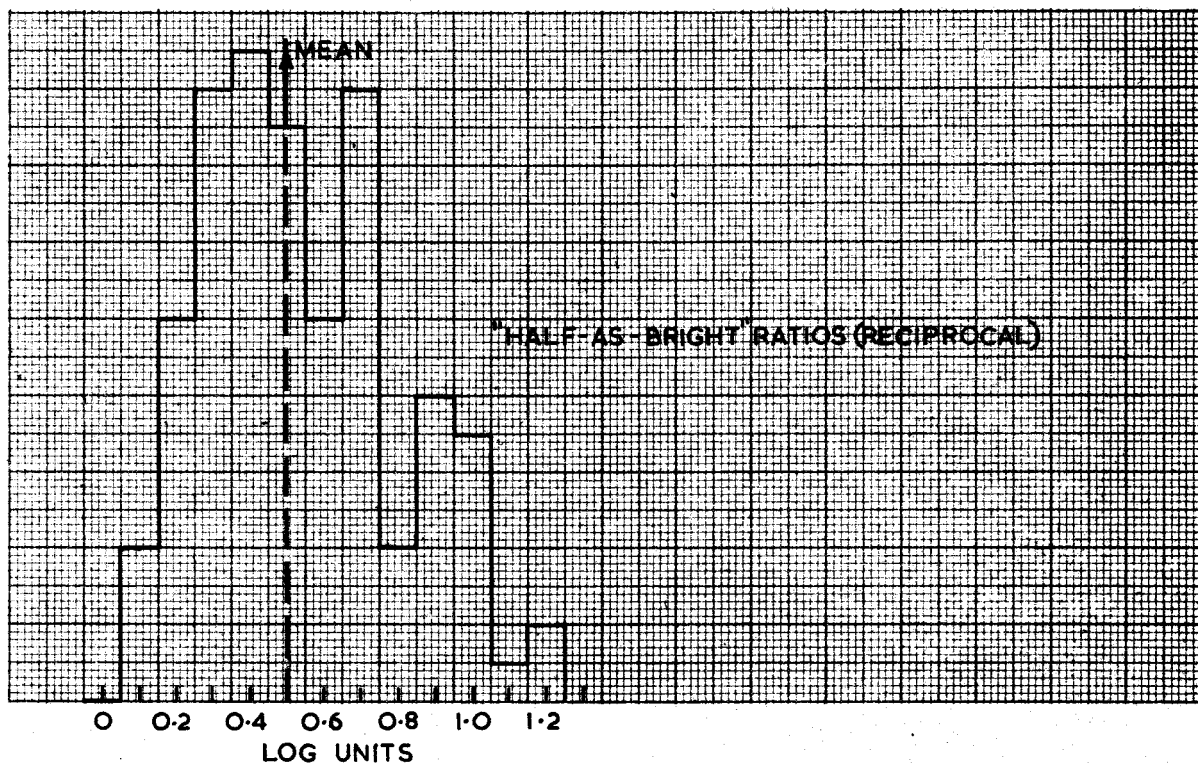


FIG. 3
DISTRIBUTION OF CHOICES OF RATIOS FOR
"TWICE-AS-BRIGHT" AND "HALF-AS-BRIGHT"

maintains his own criterion for both of the judgments, each to within less than one respective standard deviation. The ratios for the subjects do not indicate that there are two distinct groups of people, but that there is a continuous variation from a high to a low value. If the values of "twice-" and "half-as-bright" are averaged for each observer individually, (Table 3) at all brightness levels, the standard deviation of the readings for each observer is approximately three times the standard deviation of the readings made when subjects are attempting to match the two fields. If the values of "twice-" and "half-as-bright" are averaged for each brightness level individually (Table 2) for all observers, the standard deviation of the readings for each brightness level is approximately five times the standard deviation of the readings made when subjects are attempting to match the two fields.

CONCLUSIONS

If an individual is asked to make one side of a visual field twice as bright as the other side, when the surroundings are fairly dark, he can accomplish this task with a reasonable degree of consistency. He can equally well make one side half as bright as the other side. Measurements of this nature can be made by the average person and the standard deviation of repeated readings of "twice-" and "half-as-bright" will be about three times the standard deviation of repeated attempts to match the two sides of the field in brightness.

The actual value of the "two-times" ratio, however, varies greatly from one subject to another. The average for the eleven subjects used in the present experiments was about 0.56 log units; that is, approximately three and a half times the intensity for "twice-as-bright" and three-tenths of the intensity for "half-as-bright". The highest value obtained by a subject was ten times the intensity, and the lowest value by another subject was twice the intensity. These values, of course, only refer to the range of field brightnesses used in the experiment, namely 1.5 log ft. lambert to 1.5 log ft. lambert, a range of 100 to 1, but there does not seem to be any definite trend of change of value with intensity level at which a judgment is made.

On the whole these experiments differ from those in the acoustics field in that the additional intensity required to make the sensation seem "twice-as-much" was not as great, except for an extreme observer. They also differ in that it was found equally possible to measure "twice" and "half" and both ratios were of the same magnitude for each subject. It must be concluded, therefore, that such ratios, although measurable and meaningful, depend essentially on the sensation to be measured and on each individual making the measurements.

The importance of this type of measurement may be gauged by the remark of a High Court judge, quoted in the "Times" Law Reports:-

"His Lordships:- I think I can take judicial knowledge of the fact that a baby makes a noise. Presumably 25 babies make 25 times as much noise." The inaccuracy of the implied assumption that if the intensity of a noise were increased twenty-fivefold, it would sound 25 times as loud might have provided some additional weight for Defending Counsel, had he been aware of the non-linearity of such combinations.

In the case of television, it may be concluded that if it is required to make a picture appear to be twice as bright for the average viewer, then three and a half times the energy must be supplied, although some viewers will need only twice the energy. There will always be others who need ten times the energy to feel the screen is twice as bright, and it would be necessary to ascertain what percentage of people fall within each category by means of a large scale investigation, in order to decide at what level any further addition becomes uneconomic.

REFERENCES

- (1) Fletcher,
Journal Frankl. Inst., 220, p. 405, 1935.
 - (2) Robinson, D.W.
Acustica, 3, No. 5, p. 344, 1953.
 - (3) Symposium on Subjective Judgments, D.S.I.R., October 1950.
 - (4) Hopkinson, R.G., Mackenzio, R.B. and Nixon, R.D.,
Phot. Journ., B., V. 91, Jan. - Feb. 1951.
 - (5) Hopkinson, R.G.,
Nature, 170, p. 555, October 4th 1952.
-